SEQUENCE LISTING

(1) GENERAL INFORMATION:

   (i) APPLICANT: KLEIN, Michel H
            DU, Run-Pan
            EWASYSHYN, Mary E

  (ii) TITLE OF INVENTION: INFECTION DETECTION METHOD USING
       CHIMERIC PROTEIN

 (iii) NUMBER OF SEQUENCES: 38

  (iv) CORRESPONDENCE ADDRESS:
      (A) ADDRESSEE: Sim & McBurney
      (B) STREET: 6th Floor, 330 University Avenue
      (C) CITY: Toronto
      (D) STATE: Ontario
      (E) COUNTRY: Canada
      (F) ZIP: M5G 1R7

   (v) COMPUTER READABLE FORM:
      (A) MEDIUM TYPE: Floppy disk
      (B) COMPUTER: IBM PC compatible
      (C) OPERATING SYSTEM: PC-DOS/MS-DOS
      (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

  (vi) CURRENT APPLICATION DATA:
      (A) APPLICATION NUMBER:
      (B) FILING DATE:
      (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: US 08/467,961
      (B) FILING DATE: 06-JUN-1995

 (vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: US 08/001,554
      (B) FILING DATE: 06-JAN-1993

 (vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: GB 9200117.1
      (B) FILING DATE: 06-JAN-1992

 (viii) ATTORNEY/AGENT INFORMATION:
      (A) NAME: STEWART, Michael I
      (B) REGISTRATION NUMBER: 24,973
      (C) REFERENCE/DOCKET NUMBER: 1038-1000 MIS:jb

  (ix) TELECOMMUNICATION INFORMATION:
      (A) TELEPHONE: (416) 595-1155
      (B) TELEFAX: (416) 595-1163

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1844 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGTCAATAC CAACAACTAT TAGCAGTCAT ACGTGCAAGA ACAAGAAAGA AGAGATTCAA   60

AAAGCTAAAT AAGAGAAATC AAAACAAAAG GTATAGAACA CCCGAACAAC AAAATCAAAA   120

CATCCAATCC ATTTTAAACA AAAATTCCAA AAGAGACCGG CAACACAACA AGCACCAAAC   180

ACAATGCCAA CTTTAATACT GCTAATTATT ACAACAATGA TTATGGCATC TTCCTGCCAA   240

ATAGATATCA CAAAACTACA GCATGTAGGT GTATTGGTCA ACAGTCCCAA AGGGATGAAG   300

ATATCACAAA ACTTCGAAAC AAGATATCTA ATTTTGAGCC TCATACCAAA AATAGAAGAC   360

TCTAACTCTT GTGGTGACCA ACAGATCAAA CAATACAAGA GGTTATTGGA TAGACTGATC   420

ATCCCTCTAT ATGATGGATT AAGATTACAG AAAGATGTGA TAGTAACCAA TCAAGAATCC   480

AATGAAAACA CTGATCCCAG AACAAGACGA TCCTTTGGAG GGGTAATTGG AACCATTGCT   540

CTGGGAGTAG CAACCTCAGC ACAAATTACA GCGGCAGTTG CTCTGGTTGA AGCCAAGCAG   600

GCAAAATCAG ACATCGAAAA ACTCAAAGAA GCAATCAGGG ACACAAACAA AGCAGTGCAG   660

TCAGTTCAGA GCTCTATAGG AAATTTAATA GTAGCAATTA AATCAGTCCA AGATTATGTC   720

AACAACGAAA TGGTGCCATC GATTGCTAGA CTAGGTTGTG AAGCAGCAGG ACTTCAATTA   780

GGAATTGCAT TAACACAGCA TTACTCAGAA TTAACAAACA TATTTGGTGA TAACATAGGA   840

TCGTTACAAG AAAAAGGAAT AAAATTACAA GGTATAGCAT CATTATACCG CACAAATATC   900

ACAGAAATAT TCACAACATC AACAGTTGAT AAATATGATA TCTATGATCT ATTATTTACA   960

GAATCAATAA AGGTGAGAGT TATAGATGTT GATTTGAATG ATTACTCAAT CACCCTCCAA   1020

GTCAGACTCC CTTTATTAAC TAGGCTGCTG AACACTCAGA TCTACAAAGT AGATTCCATA   1080

TCATATAATA TCCAAAACAG AGAATGGTAT ATCCCTCTTC CCAGCCATAT CATGACGAAA   1140

GGGGCATTTC TAGGTGGAGC AGATGTCAAG GAATGTATAG AAGCATTCAG CAGTTATATA   1200

TGCCCTTCTG ATCCAGGATT TGTACTAAAC CATGAAATGG AGAGCTGCTT ATCAGGAAAC   1260

ATATCCCAAT GTCCAAGAAC CACGGTCACA TCAGACATTG TTCCAAGATA TGCATTTGTC   1320

AATGGAGGAG TGGTTGCAAA CTGTATAACA ACCACCTGTA CATGCAACGG AATCGACAAT   1380

```
AGAATCAATC AACCACCTGA TCAAGGAGTA AAAATTATAA CACATAAAGA ATGTAATACA      1440

ATAGGTATCA ACGGAATGCT GTTCAATACA AATAAAGAAG GAACTCTTGC ATTCTACACA      1500

CCAAATGATA TAACACTAAA TAATTCTGTT GCACTTGATC CAATTGACAT ATCAATCGAG      1560

CTTAACAAAG CCAAATCAGA TCTAGAAGAA TCAAAAGAAT GGATAAGAAG GTCAAATCAA      1620

AAACTAGATT CTATTGGAAA CTGGCATCAA TCTAGCACTA CAATCATAAT TATTTTAATA      1680

ATGATCATTA TATTGTTTAT AATTAATGTA ACGATAATTA CAATTGCAAT TAAGTATTAC      1740

AGAATTCAAA AGAGAAATCG AGTGGATCAA AATGACAAGC CATATGTACT AACAAACAAA      1800

TGACATATCT ATAGATCATT AGATATTAAA ATTATAAAAA ACTT                       1844
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 539 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Pro Thr Leu Ile Leu Leu Ile Ile Thr Thr Met Ile Met Ala Ser
1               5                   10                  15

Ser Cys Gln Ile Asp Ile Thr Lys Leu Gln His Val Gly Val Leu Val
              20                  25                  30

Asn Ser Pro Lys Gly Met Lys Ile Ser Gln Asn Phe Glu Thr Arg Tyr
            35                  40                  45

Leu Ile Leu Ser Leu Ile Pro Lys Ile Glu Asp Ser Asn Ser Cys Gly
        50                  55                  60

Asp Gln Gln Ile Lys Gln Tyr Lys Arg Leu Leu Asp Arg Leu Ile Ile
65                  70                  75                  80

Pro Leu Tyr Asp Gly Leu Arg Leu Gln Lys Asp Val Ile Val Thr Asn
                85                  90                  95

Gln Glu Ser Asn Glu Asn Thr Asp Pro Arg Thr Arg Arg Ser Phe Gly
            100                 105                 110

Gly Val Ile Gly Thr Ile Ala Leu Gly Val Ala Thr Ser Ala Gln Ile
        115                 120                 125

Thr Ala Ala Val Ala Leu Val Glu Ala Lys Gln Ala Lys Ser Asp Ile
        130                 135                 140

Glu Lys Leu Lys Glu Ala Ile Arg Asp Thr Asn Lys Ala Val Gln Ser
```

```
        145                  150                  155                  160

Val Gln Ser Ser Ile Gly Asn Leu Ile Val Ala Ile Lys Ser Val Gln
                165                  170                  175

Asp Tyr Val Asn Asn Glu Ile Val Pro Ser Ile Ala Arg Leu Gly Cys
                180                  185                  190

Glu Ala Ala Gly Leu Gln Leu Gly Ile Ala Leu Thr Gln His Tyr Ser
                195                  200                  205

Glu Leu Thr Asn Ile Phe Gly Asp Asn Ile Gly Ser Leu Gln Glu Lys
        210                  215                  220

Gly Ile Lys Leu Gln Gly Ile Ala Ser Leu Tyr Arg Thr Asn Ile Thr
225                  230                  235                  240

Glu Ile Phe Thr Thr Ser Thr Val Asp Lys Tyr Asp Ile Tyr Asp Leu
                245                  250                  255

Leu Phe Thr Glu Ser Ile Lys Val Arg Val Ile Asp Val Asp Leu Asn
                260                  265                  270

Asp Tyr Ser Ile Thr Leu Gln Val Arg Leu Pro Leu Leu Thr Arg Leu
                275                  280                  285

Leu Asn Thr Gln Ile Tyr Lys Val Asp Ser Ile Ser Tyr Asn Ile Gln
        290                  295                  300

Asn Arg Glu Trp Tyr Ile Pro Leu Pro Ser His Ile Met Thr Lys Gly
305                  310                  315                  320

Ala Phe Leu Gly Gly Ala Asp Val Lys Glu Cys Ile Glu Ala Phe Ser
                325                  330                  335

Ser Tyr Ile Cys Pro Ser Asp Pro Gly Phe Val Leu Asn His Glu Met
                340                  345                  350

Glu Ser Cys Leu Ser Gly Asn Ile Ser Gln Cys Pro Arg Thr Thr Val
        355                  360                  365

Thr Ser Asp Ile Val Pro Arg Tyr Ala Phe Val Asn Gly Gly Val Val
        370                  375                  380

Ala Asn Cys Ile Thr Thr Thr Cys Thr Cys Asn Gly Ile Asp Asn Arg
385                  390                  395                  400

Ile Asn Gln Pro Pro Asp Gln Gly Val Lys Ile Ile Thr His Lys Glu
                405                  410                  415

Cys Asn Thr Ile Gly Ile Asn Gly Met Leu Phe Asn Thr Asn Lys Glu
                420                  425                  430

Gly Thr Leu Ala Phe Tyr Thr Pro Asn Asp Ile Thr Leu Asn Asn Ser
                435                  440                  445

Val Ala Leu Asp Pro Ile Asp Ile Ser Ile Glu Leu Asn Lys Ala Lys
```

```
         450                455                460
```

Ser Asp Leu Glu Glu Ser Lys Glu Trp Ile Arg Arg Ser Asn Gln Lys
465   470   475   480

Leu Asp Ser Ile Gly Asn Trp His Gln Ser Ser Thr Thr Ile Ile Ile
                485   490   495

Ile Leu Ile Met Ile Ile Ile Leu Phe Ile Ile Asn Val Thr Ile Ile
                500   505   510

Thr Ile Ala Ile Lys Tyr Tyr Arg Ile Gln Lys Arg Asn Arg Val Asp
         515   520   525

Gln Asn Asp Lys Pro Tyr Val Leu Thr Asn Lys
    530   535

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1833 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGACAAATCC AAATTCGAGA TGGAATACTG GAAGCATACC AATCACGGAA AGGATGCTGG   60

CAATGAGCTG GAGACGTCCA TGGCTACTAA TGGCAACAAG CTCACCAATA AGATAACATA   120

TATATTATGG ACAATAATCC TGGTGTTATT ATCAATAGTC TTCATCATAG TGCTAATTAA   180

TTCCATCAAA AGTGAAAAGG CTCATGAATC ATTGCTGCAA GACATAAATA ATGAGTTTAT   240

GGAAATTACA GAAAAGATCC AAATGGCATC GGATAATACC AATGATCTAA TACAGTCAGG   300

AGTGAATACA AGGCTTCTTA CAATTCAGAG TCATGTCCAG AATTATATAC CAATATCACT   360

GACACAACAG ATGTCAGATC TTAGGAAATT CATTAGTGAA ATTACAATTA GAAATGATAA   420

TCAAGAAGTG CTGCCACAAA GAATAACACA TGATGTGGGT ATAAAACCTT TAAATCCAGA   480

TGATTTTTGG AGATGCACGT CTGGTCTTCC ATCTTTAATG AAAACTCCAA AAATAAGGTT   540

AATGCCAGGG CCGGGATTAT TAGCTATGCC AACGACTGTT GATGGCTGTA TCAGAACTCC   600

GTCCTTAGTT ATAAATGATC TGATTTATGC TTATACCTCA AATCTAATTA CTCGAGGTTG   660

TCAGGATATA GGAAAATCAT ATCAAGTCTT ACAGATAGGG ATAATAACTG TAAACTCAGA   720

CTTGGTACCT GACTTAAATC CCAGGATCTC TCATACTTTT AACATAAATG ACAATAGGAA   780

GTCATGTTCT CTAGCACTCC TAAATACAGA TGTATATCAA CTGTGTTCAA CTCCCAAAGT   840

```
TGATGAAAGA TCAGATTATG CATCATCAGG CATAGAAGAT ATTGTACTTG ATATTGTCAA      900

TTATGATGGC TCAATCTCAA CAACAAGATT TAAGAATAAT AACATAAGCT TTGATCAACC      960

TTATGCTGCA CTATACCCAT CTGTTGGACC AGGGATATAC TACAAAGGCA AAATAATATT     1020

TCTCGGGTAT GGAGGTCTTG AACATCCAAT AAATGAGAAT GTAATCTGCA ACACAACTGG     1080

GTGTCCCGGG AAAACACAGA GAGACTGCAA TCAGGCATCT CATAGTCCAT GGTTTTCAGA     1140

TAGGAGGATG GTCAACTCTA TCATTGTTGT TGACAAAGGC TTAAACTCAA TTCCAAAATT     1200

GAAGGTATGG ACGATATCTA TGAGACAGAA TTACTGGGGG TCAGAAGGAA GGTTACTTCT     1260

ACTAGGTAAC AAGATCTATA TATATACAAG ATCCACAAGT TGGCATAGCA AGTTACAATT     1320

AGGAATAATT GATATTACTG ATTACAGTGA TATAAGGATA AAATGGACAT GGCATAATGT     1380

GCTATCAAGA CCAGGAAACA ATGAATGTCC ATGGGGACAT TCATGTCCAG ATGGATGTAT     1440

AACAGGAGTA TATACTGATG CATATCCACT CAATCCCACA GGGAGCATTG TGTCATCTGT     1500

CATATTAGAT TCACAAAAAT CGAGAGTGAA CCCAGTCATA ACTTACTCAA CAGCAACCGA     1560

AAGAGTAAAC GAGCTGGCCA TCCGAAACAG AACACTCTCA GCTGGATATA CAACAACAAG     1620

CTGCATCACA CACTATAACA AAGGATATTG TTTTCATATA GTAGAAATAA ATCAGAAAAG     1680

CTTAAACACA CTTCAACCCA TGTTGTTCAA GACAGAGGTT CCAAAAAGCT GCAGTTAATC     1740

ATAATTAACC GCAATATGCA TTAACCTATC TATAATACAA GTATATGATA AGTAATCAGC     1800

AATCAGACAA TAGACAAAAG GGAAATATAA AAA                                  1833
```

(2) INFORMATION FOR SEQ ID NO:4:

  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 572 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

  Met Glu Tyr Trp Lys His Thr Asn His Gly Lys Asp Ala Gly Asn Glu
  1   5   10   15

  Leu Glu Thr Ser Met Ala Thr Asn Gly Asn Lys Leu Thr Asn Lys Ile
                  20   25   30

  Thr Tyr Ile Leu Trp Thr Ile Ile Leu Val Leu Leu Ser Ile Val Phe
                  35   40   45

Ile Ile Val Leu Ile Asn Ser Ile Lys Ser Glu Lys Ala His Glu Ser
    50   55   60

Leu Leu Gln Asp Ile Asn Asn Glu Phe Met Glu Ile Thr Glu Lys Ile
65   70   75   80

Gln Met Ala Ser Asp Asn Thr Asn Asp Leu Ile Gln Ser Gly Val Asn
                85   90   95

Thr Arg Leu Leu Thr Ile Gln Ser His Val Gln Asn Tyr Ile Pro Ile
            100   105   110

Ser Leu Thr Gln Gln Met Ser Asp Leu Arg Lys Phe Ile Ser Glu Ile
            115   120   125

Thr Ile Arg Asn Asp Asn Gln Glu Val Leu Pro Gln Arg Ile Thr His
    130   135   140

Asp Val Gly Ile Lys Pro Leu Asn Pro Asp Asp Phe Trp Arg Cys Thr
145   150   155   160

Ser Gly Leu Pro Ser Leu Met Lys Thr Pro Lys Ile Arg Leu Met Pro
                165   170   175

Gly Pro Gly Leu Leu Ala Met Pro Thr Thr Val Asp Gly Cys Ile Arg
            180   185   190

Thr Pro Ser Leu Val Ile Asn Asp Leu Ile Tyr Ala Tyr Thr Ser Asn
        195   200   205

Leu Ile Thr Arg Gly Cys Gln Asp Ile Gly Lys Ser Tyr Gln Val Leu
    210   215   220

Gln Ile Gly Ile Ile Thr Val Asn Ser Asp Leu Val Pro Asp Leu Asn
225   230   235   240

Pro Arg Ile Ser His Thr Phe Asn Ile Asn Asp Asn Arg Lys Ser Cys
                245   250   255

Ser Leu Ala Leu Leu Asn Thr Asp Val Tyr Gln Leu Cys Ser Thr Pro
            260   265   270

Lys Val Asp Glu Arg Ser Asp Tyr Ala Ser Ser Gly Ile Glu Asp Ile
        275   280   285

Val Leu Asp Ile Val Asn Tyr Asp Gly Ser Ile Ser Thr Thr Arg Phe
    290   295   300

Lys Asn Asn Asn Ile Ser Phe Asp Gln Pro Tyr Ala Ala Leu Tyr Pro
305   310   315   320

Ser Val Gly Pro Gly Ile Tyr Tyr Lys Gly Lys Ile Ile Phe Leu Gly
                325   330   335

Tyr Gly Gly Leu Glu His Pro Ile Asn Glu Asn Val Ile Cys Asn Thr
            340   345   350

```
Thr Gly Cys Pro Gly Lys Thr Gln Arg Asp Cys Asn Gln Ala Ser His
        355                 360                 365

Ser Pro Trp Phe Ser Asp Arg Arg Met Val Asn Ser Ile Ile Val Val
    370                 375                 380

Asp Lys Gly Leu Asn Ser Ile Pro Lys Leu Lys Val Trp Thr Ile Ser
385                 390                 395                 400

Met Arg Gln Asn Tyr Trp Gly Ser Glu Gly Arg Leu Leu Leu Leu Gly
                405                 410                 415

Asn Lys Ile Tyr Ile Tyr Thr Arg Ser Thr Ser Trp His Ser Lys Leu
            420                 425                 430

Gln Leu Gly Ile Ile Asp Ile Thr Asp Tyr Ser Asp Ile Arg Ile Lys
        435                 440                 445

Trp Thr Trp His Asn Val Leu Ser Arg Pro Gly Asn Asn Glu Cys Pro
    450                 455                 460

Trp Gly His Ser Cys Pro Asp Gly Cys Ile Thr Gly Val Tyr Thr Asp
465                 470                 475                 480

Ala Tyr Pro Leu Asn Pro Thr Gly Ser Ile Val Ser Ser Val Ile Leu
                485                 490                 495

Asp Ser Gln Lys Ser Arg Val Asn Pro Val Ile Thr Tyr Ser Thr Ala
            500                 505                 510

Thr Glu Arg Val Asn Glu Leu Ala Ile Arg Asn Arg Thr Leu Ser Ala
        515                 520                 525

Gly Tyr Thr Thr Thr Ser Cys Ile Thr His Tyr Asn Lys Gly Tyr Cys
    530                 535                 540

Phe His Ile Val Glu Ile Asn Gln Lys Ser Leu Asn Thr Leu Gln Pro
545                 550                 555                 560

Met Leu Phe Lys Thr Glu Val Pro Lys Ser Cys Ser
                565                 570
```

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1886 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
ATGGAGTTGC CAATCCTCAA AGCAAATGCA ATTACCACAA TCCTCGCTGC AGTCACATTT          60
```

```
TGCTTTGCTT CTAGTCAAAA CATCACTGAA GAATTTTATC AATCAACATG CAGTGCAGTT     120

AGCAAAGGCT ATCTTAGTGC TCTAAGAACT GGTTGGTATA CTAGTGTTAT AACTATAGAA     180

TTAAGTAATA TCAAGGAAAA TAAGTGTAAT GGAACAGATG CTAAGGTAAA ATTGATGAAA     240

CAAGAATTAG ATAAATATAA AAATGCTGTA ACAGAATTGC AGTTGCTCAT GCAAAGCACA     300

CCAGCAGCAA ACAATCGAGC CAGAAGAGAA CTACCAAGGT TTATGAATTA TACACTCAAC     360

AATACCAAAA AAACCAATGT AACATTAAGC AAGAAAGGA AAAGAAGATT TCTTGGTTTT     420

TTGTTAGGTG TTGGATCTGC AATCGCCAGT GGCATTGCTG TATCTAAGGT CCTGCACTTA     480

GAAGGAGAAG TGAACAAGAT CAAAAGTGCT CTACTATCCA CAAACAAGGC CGTAGTCAGC     540

TTATCAAATG GAGTTAGTGT CTTAACCAGC AAAGTGTTAG ACCTCAAAAA CTATATAGAT     600

AAACAATTGT TACCTATTGT GAATAAGCAA AGCTGCAGAA TATCAAATAT AGAAACTGTG     660

ATAGAGTTCC AACAAAGAA CAACAGACTA CTAGAGATTA CCAGGGAATT TAGTGTTAAT     720

GCAGGTGTAA CTACACCTGT AAGCACTTAC ATGTTAACTA ATAGTGAATT ATTGTCATTA     780

ATCAATGATA TGCCTATAAC AAATGATCAG AAAAAGTTAA TGTCCAACAA TGTTCAAATA     840

GTTAGACAGC AAAGTTACTC TATCATGTCC ATAATAAAG AGGAAGTCTT AGCATATGTA     900

GTACAATTAC CACTATATGG TGTGATAGAT ACACCTTGTT GGAAATTACA CACATCCCCT     960

CTATGTACAA CCAACACAAA AGAAGGGTCA AACATCTGTT TAACAAGAAC TGACAGAGGA    1020

TGGTACTGTG ACAATGCAGG ATCAGTATCT TTCTTCCCAC AAGCTGAAAC ATGTAAAGTT    1080

CAATCGAATC GAGTATTTTG TGACACAATG AACAGTTTAA CATTACCAAG TGAAGTAAAT    1140

CTCTGCAATG TTGACATATT CAATCCCAAA TATGATTGTA AAATTATGAC TTCAAAAACA    1200

GATGTAAGCA GCTCCGTTAT CACATCTCTA GGAGCCATTG TGTCATGCTA TGGCAAAACT    1260

AAATGTACAG CATCCAATAA AAATCGTGGA ATCATAAAGA CATTTTCTAA CGGGTGTGAT    1320

TATGTATCAA ATAAAGGGGT GGACACTGTG TCTGTAGGTA ACACATTATA TTATGTAAAT    1380

AAGCAAGAAG GCAAAAGTCT CTATGTAAAA GGTGAACCAA TAATAAATTT CTATGACCCA    1440

TTAGTATTCC CCTCTGATGA ATTTGATGCA TCAATATCTC AAGTCAATGA GAAGATTAAC    1500

CAGAGTTTAG CATTTATTCG TAAATCCGAT GAATTATTAC ATAATGTAAA TGCTGGTAAA    1560

TCAACCACAA ATATCATGAT AACTACTATA ATTATAGTGA TTATAGTAAT ATTGTTATCA    1620

TTAATTGCTG TTGGACTGCT CCTATACTGT AAGGCCAGAA GCACACCAGT CACACTAAGC    1680

AAGGATCAAC TGAGTGGTAT AAATAATATT GCATTTAGTA ACTGAATAAA AATAGCACCT    1740
```

AATCATGTTC TTACAATGGT TTACTATCTG CTCATAGACA ACCCATCTAT CATTGGATTT   1800

TCTTAAAATC TGAACTTCAT CGAAACTCTT ATCTATAAAC CATCTCACTT ACACTATTTA   1860

AGTAGATTCC TAGTTTATAG TTATAT   1886

(2) INFORMATION FOR SEQ ID NO:6:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 594 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

    Met Glu Leu Pro Ile Leu Lys Ala Asn Ala Ile Thr Thr Ile Leu Ala
    1   5   10   15

    Ala Val Thr Phe Cys Phe Ala Ser Ser Gln Asn Ile Thr Glu Glu Phe
                    20   25   30

    Tyr Gln Ser Thr Cys Ser Ala Val Ser Lys Gly Tyr Leu Ser Ala Leu
                35   40   45

    Arg Thr Gly Trp Tyr Thr Ser Val Ile Thr Ile Glu Leu Ser Asn Ile
        50   55   60

    Lys Glu Asn Lys Cys Asn Gly Thr Asp Ala Lys Val Lys Leu Met Lys
    65   70   75   80

    Gln Glu Leu Asp Lys Tyr Lys Asn Ala Val Thr Glu Leu Gln Leu Leu
                    85   90   95

    Met Gln Ser Thr Pro Ala Ala Asn Asn Arg Ala Arg Arg Glu Leu Pro
                    100   105   110

    Arg Phe Met Asn Tyr Thr Leu Asn Asn Thr Lys Lys Thr Asn Val Thr
                115   120   125

    Leu Ser Lys Lys Arg Lys Arg Arg Phe Leu Gly Phe Leu Leu Gly Val
        130   135   140

    Gly Ser Ala Ile Ala Ser Gly Ile Ala Val Ser Lys Val Leu His Leu
    145   150   155   160

    Glu Gly Glu Val Asn Lys Ile Lys Ser Ala Leu Leu Ser Thr Asn Lys
                    165   170   175

    Ala Val Val Ser Leu Ser Asn Gly Val Ser Val Leu Thr Ser Lys Val
                    180   185   190

    Leu Asp Leu Lys Asn Tyr Ile Asp Lys Gln Leu Leu Pro Ile Val Asn
                195   200   205

```
Lys Gln Ser Cys Arg Ile Ser Asn Ile Glu Thr Val Ile Glu Phe Gln
    210                 215                 220

His Lys Asn Asn Arg Leu Leu Glu Ile Thr Arg Glu Phe Ser Val Asn
225                 230                 235                 240

Ala Gly Val Thr Thr Pro Val Ser Thr Tyr Met Leu Thr Asn Ser Glu
                245                 250                 255

Leu Leu Ser Leu Ile Asn Asp Met Pro Ile Thr Asn Asp Gln Lys Lys
                260                 265                 270

Leu Met Ser Asn Asn Val Gln Ile Val Arg Gln Gln Ser Tyr Ser Ile
                275                 280                 285

Met Ser Ile Ile Lys Glu Glu Val Leu Ala Tyr Val Val Gln Leu Pro
    290                 295                 300

Leu Tyr Gly Val Ile Asp Thr Pro Cys Trp Lys Leu His Thr Ser Pro
305                 310                 315                 320

Leu Cys Thr Thr Asn Thr Lys Glu Gly Ser Asn Ile Cys Leu Thr Arg
                325                 330                 335

Thr Asp Arg Gly Trp Tyr Cys Asp Asn Ala Gly Ser Val Ser Phe Phe
                340                 345                 350

Pro Gln Ala Glu Thr Cys Lys Val Gln Ser Asn Arg Val Phe Cys Asp
    355                 360                 365

Thr Met Asn Ser Leu Thr Leu Pro Ser Glu Val Asn Leu Cys Asn Val
    370                 375                 380

Asp Ile Phe Asn Pro Lys Tyr Asp Cys Lys Ile Met Thr Ser Lys Thr
385                 390                 395                 400

Asp Val Ser Ser Ser Val Ile Thr Ser Leu Gly Ala Ile Val Ser Cys
                405                 410                 415

Tyr Gly Lys Thr Lys Cys Thr Ala Ser Asn Lys Asn Arg Gly Ile Ile
                420                 425                 430

Lys Thr Phe Ser Asn Gly Cys Asp Tyr Val Ser Asn Lys Gly Val Asp
                435                 440                 445

Thr Val Ser Val Gly Asn Thr Leu Tyr Tyr Val Asn Lys Gln Glu Gly
    450                 455                 460

Lys Ser Leu Tyr Val Lys Gly Glu Pro Ile Ile Asn Phe Tyr Asp Pro
465                 470                 475                 480

Leu Val Phe Pro Ser Asp Glu Phe Asp Ala Ser Ile Ser Gln Val Asn
                485                 490                 495

Glu Lys Ile Asn Leu Val Phe Pro Ser Asp Glu Phe Asp Ala Ser Ile
                500                 505                 510
```

```
        Ser Gln Val Asn Glu Lys Ile Asn Gln Ser Leu Ala Phe Ile Arg Lys
                515                 520                 525

        Ser Asp Glu Leu Leu His Asn Val Asn Ala Gly Lys Ser Thr Thr Asn
            530                 535                 540

        Ile Met Ile Thr Thr Ile Ile Ile Glu Ile Ile Val Ile Leu Leu Ser
        545                 550                 555                 560

        Leu Ile Ala Val Gly Leu Leu Leu Tyr Cys Lys Ala Arg Ser Thr Pro
                    565                 570                 575


        Val Thr Leu Ser Lys Asp Gln Leu Ser Gly Ile Asn Asn Ile Ala Phe
                580                 585                 590

        Ser Asn



(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 920 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGCAAACATG TCCAAAAACA AGGACCAACG CACCGCTAAG ACACTAGAAA AGACCTGGGA        60

CACTCTCAAT CATTTATTAT TCATATCATC GGGCTTATAT AAGTTAAATC TTAAATCTGT       120

AGCACAAATC ACATTATCCA TTCTGGCAAT GATAATCTCA ACTTCACTTA TAATTACAGC       180

CATCATATTC ATAGCCTCGG CAAACCACAA AGTCACACTA CAACTGCAA TCATACAAGA        240

TGCAACAAGC CAGATCAAGA ACACAACCCC AACATACCTC ACTCAGGATC CTCAGCTTGG       300

AATCAGCTTC TCCAATCTGT CTGAAATTAC ATCACAAACC ACCACCATAC TAGCTTCAAC       360

AACACCAGGA GTCAAGTCAA ACCTGCAACC CACAACAGTC AAGACTAAAA ACACAACAAC       420

AACCCAAACA CAACCCAGCA AGCCCACTAC AAAACAACGC CAAAACAAAC CACCAAACAA       480

ACCCAATAAT GATTTTCACT TCGAAGTGTT TAACTTTGTA CCCTGCAGCA TATGCAGCAA       540

CAATCCAACC TGCTGGGCTA TCTGCAAAAG AATACCAAAC AAAAAACCAG GAAAGAAAAC       600

CACCACCAAG CCTACAAAAA AACCAACCTT CAAGACAACC AAAAAGATC TCAAACCTCA        660

AACCACTAAA CCAAAGGAAG TACCCACCAC CAAGCCCACA GAAGAGCCAA CCATCAACAC       720
```

CACCAAAACA AACATCACAA CTACACTGCT CACCAACAAC ACCACAGGAA ATCCAAAACT   780

CACAAGTCAA ATGGAAACCT TCCACTCAAC CTCCTCCGAA GGCAATCTAA GCCCTTCTCA   840

AGTCTCCACA ACATCCGAGC ACCCATCACA ACCCTCATCT CCACCCAACA CAACACGCCA   900

GTAGTTATTA AAAAAAAAAA   920

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 298 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

    Met Ser Lys Asn Lys Asp Gln Arg Thr Ala Lys Thr Leu Glu Lys Thr
    1   5   10   15

    Trp Asp Thr Leu Asn His Leu Leu Phe Ile Ser Ser Gly Leu Tyr Lys
                20   25   30

    Leu Asn Leu Lys Ser Val Ala Gln Ile Thr Leu Ser Ile Leu Ala Met
                35   40   45

    Ile Ile Ser Thr Ser Leu Ile Ile Thr Ala Ile Ile Phe Ile Ala Ser
        50   55   60

    Ala Asn His Lys Val Thr Leu Thr Thr Ala Ile Ile Gln Asp Ala Thr
    65   70   75   80

    Ser Gln Ile Lys Asn Thr Thr Pro Thr Tyr Leu Thr Gln Asp Pro Gln
                    85   90   95

    Leu Gly Ile Ser Phe Ser Asn Leu Ser Glu Ile Thr Ser Gln Thr Thr
                    100   105   110

    Thr Ile Leu Ala Ser Thr Thr Pro Gly Val Lys Ser Asn Leu Gln Pro
            115   120   125

    Thr Thr Val Lys Thr Lys Asn Thr Thr Thr Thr Gln Thr Gln Pro Ser
        130   135   140

    Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn Lys Pro Asn
    145   150   155   160

    Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
                    165   170   175

    Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys
                180   185   190

```
Lys Pro Gly Lys Lys Thr Thr Thr Lys Pro Thr Lys Lys Pro Thr Phe
        195                 200             205

Lys Thr Thr Lys Lys Asp Leu Lys Pro Gln Thr Thr Lys Pro Lys Glu
        210                 215             220

Val Pro Thr Thr Lys Pro Thr Glu Glu Pro Thr Ile Asn Thr Thr Lys
225                 230             235                 240

Thr Asn Ile Thr Thr Thr Leu Leu Thr Asn Asn Thr Thr Gly Asn Pro
                245             250                 255

Lys Leu Thr Ser Gln Met Glu Thr Phe His Ser Thr Ser Ser Glu Gly
        260             265                 270

Asn Leu Ser Pro Ser Gln Val Ser Thr Thr Ser Glu His Pro Ser Gln
        275             280                 285

Pro Ser Ser Pro Pro Asn Thr Thr Arg Gln
    290             295
```

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATCAATCAAA GGTCCTGTGA TAATAG   26

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATGACTTGA TAATGAG   17

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 86 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AATTCATGGA GTTGCTAATC CTCAAAGCAA ATGCAATTAC CACAATCCTC ACTGCAGTCA   60

CATTTTGTTT TGCTTCTGGT TCTAAG   86

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACTGGCATCA ATCTAGCACT ACATGAG   27

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 136 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTCATGCC AACTTTAATA CTGCTAATTA TTACAACAAT GATTATGGCA TCTTCCTGCC   60

AAATAGATAT CACAAAACTA CAGCATGTAG GTGTATTGGT CAACAGTCCC AAAGGGATGA   120

AGATATCACA AAACTT   136

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 94 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATCATGGAGA TAATTAAAAT GATAACCATC TCGCAAATAA ATAAGTATTT TACTGTTTTC   60

GTAACAGTTT TGTAATAAAA AAACCTATAA ATAG   94

(2) INFORMATION FOR SEQ ID NO:15:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 141 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATCATGGAGA TAATTAAAAT GATAACCATC TCGCAAATAA ATAAGTATTT TACTGTTTTC   60

GTAACAGTTT TGTAATAAAA AAACCTATAA ATATTCCGGA ATTCAGATCT GCAGCGGCCG   120

CTCCATCTAG AAGGTACCCG G   141

(2) INFORMATION FOR SEQ ID NO:16:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 31 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CATGACTAAT TCCATCAAAA GTGAAAAGGC T   31

(2) INFORMATION FOR SEQ ID NO:17:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 18 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAAGAAAAAG GAATAAAA   18

(2) INFORMATION FOR SEQ ID NO:18:

     (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATTTCTGTGA TATTTGTGCG GTATAATGAT GCTATACCT   39

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 18 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CAGGAGAAGG GTATCAAG   18

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AGGAGAAGGG TATCAAG   17

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 94 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATCATGGAGA TAATTAAAAT GATAACCATC TCGCAAATAA ATAAGTATTT TACTGTTTTC   60

GTAACAGTTT TGTAATAAAA AAACCTATAA ATAG

(2) INFORMATION FOR SEQ ID NO:22:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 6 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

      Gln Glu Lys Gly Ile Lys
        1   5

(2) INFORMATION FOR SEQ ID NO:23:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 6 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

      Gln Glu Lys Gly Ile Lys
        1   5

(2) INFORMATION FOR SEQ ID NO:24:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 20 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATCAATCTAG CACTACACAG   20

(2) INFORMATION FOR SEQ ID NO:25:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 1617 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
ATGCCAACTT TAATACTGCT AATTATTACA ACAATGATTA TGGCATCTTC CTGCCAAATA        60

GATATCACAA AACTACAGCA TGTAGGTGTA TTGGTCAACA GTCCCAAAGG GATGAAGATA       120

TCACAAAACT TCGAAACAAG ATATCTAATT TTGAGCCTCA TACCAAAAAT AGAAGACTCT       180

AACTCTTGTG GTGACCAACA GATCAAACAA TACAAGAGGT TATTGGATAG ACTGATCATC       240

CCTCTATATG ATGGATTAAG ATTACAGAAA GATGTGATAG TAACCAATCA AGAATCCAAT       300

GAAAACACTG ATCCCAGAAC AAGACGATCC TTTGGAGGGG TAATTGGAAC CATTGCTCTG       360

GGAGTAGCAA CCTCAGCACA AATTACAGCG GCAGTTGCTC TGGTTGAAGC CAAGCAGGCA       420

AAATCAGACA TCGAAAAACT CAAAGAAGCA ATCAGGGACA CAAACAAAGC AGTGCAGTCA       480

GTTCAGAGCT CTATAGGAAA TTTAATAGTA GCAATTAAAT CAGTCCAAGA TTATGTCAAC       540

AACGAAATGG TGCCATCGAT TGCTAGACTA GGTTGTGAAG CAGCAGGACT TCAATTAGGA       600

ATTGCATTAA CACAGCATTA CTCAGAATTA ACAAACATAT TTGGTGATAA CATAGGATCG       660

TTACAAGAAA AAGGAATAAA ATTACAAGGT ATAGCATCAT TATACCGCAC AAATATCACA       720

GAAATATTCA CAACATCAAC AGTTGATAAA TATGATATCT ATGATCTATT ATTTACAGAA       780

TCAATAAAGG TGAGAGTTAT AGATGTTGAT TTGAATGATT ACTCAATCAC CCTCCAAGTC       840

AGACTCCCTT TATTAACTAG GCTGCTGAAC ACTCAGATCT ACAAAGTAGA TTCCATATCA       900

TATAATATCC AAAACAGAGA ATGGTATATC CCTCTTCCCA GCCATATCAT GACGAAAGGG       960

GCATTTCTAG GTGGAGCAGA TGTCAAGGAA TGTATAGAAG CATTCAGCAG TTATATATGC      1020

CCTTCTGATC CAGGATTTGT ACTAAACCAT GAAATGGAGA GCTGCTTATC AGGAAACATA      1080

TCCCAATGTC CAAGAACCAC GGTCACATCA GACATTGTTC CAGATATGC ATTTGTCAAT      1140

GGAGGAGTGG TTGCAAACTG TATAACAACC ACCTGTACAT GCAACGGAAT CGACAATAGA      1200

ATCAATCAAC CACCTGATCA AGGAGTAAAA ATTATAACAC ATAAAGAATG TAATACAATA      1260

GGTATCAACG GAATGCTGTT CAATACAAAT AAAGAAGGAA CTCTTGCATT CTACACACCA      1320

AATGATATAA CACTAAATAA TTCTGTTGCA CTTGATCCAA TTGACATATC AATCGAGCTT      1380

AACAAAGCCA AATCAGATCT AGAAGAATCA AAAGAATGGA TAAGAAGGTC AAATCAAAAA      1440

CTAGATTCTA TTGGAAACTG GCATCAATCT AGCACTACAA TCATAATTAT TTTAATAATG      1500

ATCATTATAT TGTTTATAAT TAATGTAACG ATAATTACAA TTGCAATTAA GTATTACAGA      1560

ATTCAAAAGA GAAATCGAGT GGATCAAAAT GACAAGCCAT ATGTACTAAC AAACAAA        1617
```

(2) INFORMATION FOR SEQ ID NO:26:

```
    (i)  SEQUENCE CHARACTERISTICS:
         (A)  LENGTH: 1715 base pairs
         (B)  TYPE: nucleic acid
         (C)  STRANDEDNESS: single
         (D)  TOPOLOGY: linear



    (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATGGAATACT GGAAGCATAC CAATCACGGA AAGGATGCTG GCAATGAGCT GGAGACGTCC          60

ATGGCTACTA ATGGCAACAA GCTCACCAAT AAGATAACAT ATATATTATG GACAATAATC         120

CTGGTGTTAT TATCAATAGT CTTCATCATA GTGCTAATTA ATTCCATCAA AAGTGAAAAG         180

GCTCATGAAT CATTGCTGCA AGACATAAAT AATGAGTTTA TGGAAATTAC AGAAAAGATC         240

CAAATGGCAT CGGATAATAC CAATGATCTA ATACAGTCAG GAGTGAATAC AAGGCTTCTT         300

ACAATTCAGA GTCATGTCCA GAATTATATA CCAATATCAC TGACACAACA GATGTCAGAT         360

CTTAGGAAAT TCATTAGTGA AATTACAATT AGAAATGATA ATCAAGAAGT GCTGCCACAA         420

AGAATAACAC ATGATGTGGG TATAAAACCT TTAAATCCAG ATGATTTTTG GAGATGCACG         480

TCTGGTCTTC CATCTTTAAT GAAAACTCCA AAAATAAGGT TAATGCCAGG GCCGGGATTA         540

TTAGCTATGC CAACGACTGT TGATGGCTGT ATCAGAACTC CGTCCTTAGT TATAAATGAT         600

CTGATTTATG CTTATACCTC AAATCTAATT ACTCGAGGTT GTCAGGATAT AGGAAAATCA         660

TATCAAGTCT TACAGATAGG GATAATAACT GTAAACTCAG ACTTGGTACC TGACTTAAAT         720

CCCAGGATCT CTCATACTTT TAACATAAAT GACAATAGGA AGTCATGTTC TCTAGCACTC         780

CTAAATACAG ATGTATATCA ACTGTGTTCA ACTCCCAAAG TTGATGAAAG ATCAGATTAT         840

GCATCATCAG GCATAGAAGA TATTGTACTT GATATTGTCA ATTATGATGG CTCAATCTCA         900

ACAACAAGAT TTAAGAATAA TAACATAAGC TTTGATCAAC CTTATGCTGC ACTATACCCA         960

TCTGTTGGAC CAGGGATATA CTACAAAGGC AAAATAATAT TTCTCGGGTA TGGAGGTCTT        1020

GAACATCCAA TAAATGAGAA TGTAATCTGC AACACAACTG GGTGTCCCGG GAAAACACAG        1080

AGAGACTGCA ATCAGGCATC TCATAGTCCA TGGTTTTCAG ATAGGAGGAT GGTCAACTCT        1140

ATCATTGTTG TTGACAAAGG CTTAAACTCA ATTCCAAAAT TGAAGGTATG GACGATATCT        1200

ATGAGACAGA ATTACTGGGG GTCAGAAGGA AGGTTACTTC TACTAGGTAA CAAGATCTAT        1260

ATATATACAA GATCCACAAG TTGGCATAGC AAGTTACAAT TAGGAATAAT TGATATTACT        1320

GATTACAGTG ATATAAGGAT AAAATGGACA TGGCATAATG TGCTATCAAG ACCAGGAAAC        1380
```

```
AATGAATGTC CATGGGGACA TTCATGTCCA GATGGATGTA TAACAGGAGT ATATACTGAT      1440

GCATATCCAC TCAATCCCAC AGGGAGCATT GTGTCATCTG TCATATTAGA TTCACAAAAA      1500

TCGAGAGTGA ACCCAGTCAT AACTTACTCA ACAGCAACCG AAAGAGTAAA CGAGCTGGCC      1560

ATCCGAAACA GAACACTCTC AGCTGGATAT ACAACAACAA GCTGCATCAC ACACTATAAC      1620

AAAGGATATT GTTTTCATAT AGTAGAAATA AATCAGAAAA GCTTAAACAC ACTTCAACCC      1680

ATGTTGTTCA AGACAGAGGT TCCAAAAAGC TGCAG                                1715
```

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 1722 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
ATGGAGTTGC CAATCCTCAA AGCAAATGCA ATTACCACAA TCCTCGCTGC AGTCACATTT        60

TGCTTTGCTT CTAGTCAAAA CATCACTGAA GAATTTATC AATCAACATG CAGTGCAGTT       120

AGCAAAGGCT ATCTTAGTGC TCTAAGAACT GGTTGGTATA CTAGTGTTAT AACTATAGAA       180

TTAAGTAATA TCAAGGAAAA TAAGTGTAAT GGAACAGATG CTAAGGTAAA ATTGATGAAA       240

CAAGAATTAG ATAAATATAA AAATGCTGTA ACAGAATTGC AGTTGCTCAT GCAAAGCACA       300

CCAGCAGCAA ACAATCGAGC CAGAAGAGAA CTACCAAGGT TTATGAATTA TACACTCAAC       360

AATACCAAAA AAACCAATGT AACATTAAGC AAGAAAAGGA AAGAAGATT TCTTGGTTTT        420

TTGTTAGGTG TTGGATCTGC AATCGCCAGT GGCATTGCTG TATCTAAGGT CCTGCACTTA       480

GAAGGAGAAG TGAACAAGAT CAAAAGTGCT CTACTATCCA CAAACAAGGC CGTAGTCAGC       540

TTATCAAATG GAGTTAGTGT CTTAACCAGC AAAGTGTTAG ACCTCAAAAA CTATATAGAT       600

AAACAATTGT TACCTATTGT GAATAAGCAA AGCTGCAGAA TATCAAATAT AGAAACTGTG       660

ATAGAGTTCC AACAAAAGAA CAACAGACTA CTAGAGATTA CCAGGGAATT TAGTGTTAAT       720

GCAGGTGTAA CTACACCTGT AAGCACTTAC ATGTTAACTA ATAGTGAATT ATTGTCATTA       780

ATCAATGATA TGCCTATAAC AAATGATCAG AAAAAGTTAA TGTCCAACAA TGTTCAAATA       840

GTTAGACAGC AAAGTTACTC TATCATGTCC ATAATAAAAG AGGAAGTCTT AGCATATGTA       900

GTACAATTAC CACTATATGG TGTGATAGAT ACACCTTGTT GGAAATTACA CACATCCCCT       960
```

CTATGTACAA CCAACACAAA AGAAGGGTCA AACATCTGTT TAACAAGAAC TGACAGAGGA   1020

TGGTACTGTG ACAATGCAGG ATCAGTATCT TTCTTCCCAC AAGCTGAAAC ATGTAAAGTT   1080

CAATCGAATC GAGTATTTTG TGACACAATG AACAGTTTAA CATTACCAAG TGAAGTAAAT   1140

CTCTGCAATG TTGACATATT CAATCCCAAA TATGATTGTA AAATTATGAC TTCAAAAACA   1200

GATGTAAGCA GCTCCGTTAT CACATCTCTA GGAGCCATTG TGTCATGCTA TGGCAAAACT   1260

AAATGTACAG CATCCAATAA AAATCGTGGA ATCATAAAGA CATTTTCTAA CGGGTGTGAT   1320

TATGTATCAA ATAAAGGGGT GGACACTGTG TCTGTAGGTA ACACATTATA TTATGTAAAT   1380

AAGCAAGAAG GCAAAAGTCT CTATGTAAAA GGTGAACCAA TAATAAATTT CTATGACCCA   1440

TTAGTATTCC CCTCTGATGA ATTTGATGCA TCAATATCTC AAGTCAATGA GAAGATTAAC   1500

CAGAGTTTAG CATTTATTCG TAAATCCGAT GAATTATTAC ATAATGTAAA TGCTGGTAAA   1560

TCAACCACAA ATATCATGAT AACTACTATA ATTATAGTGA TTATAGTAAT ATTGTTATCA   1620

TTAATTGCTG TTGGACTGCT CCTATACTGT AAGGCCAGAA GCACACCAGT CACACTAAGC   1680

AAGGATCAAC TGAGTGGTAT AAATAATATT GCATTTAGTA AC   1722

(2) INFORMATION FOR SEQ ID NO:28:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 894 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATGTCCAAAA ACAAGGACCA ACGCACCGCT AAGACACTAG AAAAGACCTG GGACACTCTC   60

AATCATTTAT TATTCATATC ATCGGGCTTA TATAAGTTAA ATCTTAAATC TGTAGCACAA   120

ATCACATTAT CCATTCTGGC AATGATAATC TCAACTTCAC TTATAATTAC AGCCATCATA   180

TTCATAGCCT CGGCAAACCA CAAAGTCACA CTAACAACTG CAATCATACA AGATGCAACA   240

AGCCAGATCA AGAACACAAC CCCAACATAC CTCACTCAGG ATCCTCAGCT TGGAATCAGC   300

TTCTCCAATC TGTCTGAAAT TACATCACAA ACCACCACCA TACTAGCTTC AACAACACCA   360

GGAGTCAAGT CAAACCTGCA ACCCACAACA GTCAAGACTA AAAACACAAC AACAACCCAA   420

ACACAACCCA GCAAGCCCAC TACAAAACAA CGCCAAAACA AACCACCAAA CAAACCCAAT   480

AATGATTTTC ACTTCGAAGT GTTTAACTTT GTACCCTGCA GCATATGCAG CAACAATCCA   540

ACCTGCTGGG CTATCTGCAA AAGAATACCA AACAAAAAAC CAGGAAAGAA AACCACCACC   600

AAGCCTACAA AAAAACCAAC CTTCAAGACA ACCAAAAAAG ATCTCAAACC TCAAACCACT   660

AAACCAAAGG AAGTACCCAC CACCAAGCCC ACAGAAGAGC CAACCATCAA CACCACCAAA   720

ACAAACATCA CAACTACACT GCTCACCAAC AACACCACAG GAAATCCAAA ACTCACAAGT   780

CAAATGGAAA CCTTCCACTC AACCTCCTCC GAAGGCAATC TAAGCCCTTC TCAAGTCTCC   840

ACAACATCCG AGCACCCATC ACAACCCTCA TCTCCACCCA ACACAACACG CCAG   894

    (2) INFORMATION FOR SEQ ID NO:29:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 32 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGTAGTTAGT TTCCAGGACA CTATTATCCT AG   32

    (2) INFORMATION FOR SEQ ID NO:30:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 17 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TGAACTATTA CTCCTAG   17

    (2) INFORMATION FOR SEQ ID NO:31:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 85 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTACCTCAAC GATTAGGAGT TTCGTTTACG TTAATGGTGT TAGGAGTGAC GTCAGTGTAA   60

AACAAAACGA AGACCAAGAT TCCAG

(2) INFORMATION FOR SEQ ID NO:32:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CGTAGTTAGA TCGTGATGTA CTCCTAG   27

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 136 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GTACGGTTGA AATTATGACG ATTAATAATG TTGTTACTAA TACCGTAGAA GGACGGTTTA   60

TCTATAGTGT TTTGATGTCG TACATCCACA TAACCAGTTG TCAGGGTTTC CCTACTTCTA   120

TAGTGTTTTG AAGCTT   136

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 98 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TAGTACCTCT ATTAATTTTA CTATTGGTAG AGCGTTTATT TATTCATAAA ATGACAAAAG   60

CATTGTCAAA ACATTATTTT TTTGGATATT TATCTTAA   98

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 145 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TAGTACCTCT ATTAATTTTA CTATTGGTAG AGCGTTTATT TATTCATAAA ATGACAAAAG   60

CATTGTCAAA ACATTATTTT TTTGGATATT TATAAGGCCT TAAGTCTAGA CGTCGCCGGC   120

GAGGTAGATC TTCCATGGGC CCTAG   145

(2) INFORMATION FOR SEQ ID NO:36:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 31 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TGATTAAGGT AGTTTTCACT TTTCCGAGTA C   31

(2) INFORMATION FOR SEQ ID NO:37:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 39 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TAAAGACACT ATAAACACGC CATATTACTA CGATATGGA   39

(2) INFORMATION FOR SEQ ID NO:38:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 10 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AGGACAAAAG